Cassandra Crossing 675- Le promesse dei tech-bro in un bagno di realtà

Photo by Sung Jin Cho on Unsplash

(675) - L'escalation della propaganda dei tech-bro ha raggiunto un nuovo livello, reso necessario per non far sgonfiare la bolla. Le persone normali necessitano di un bagno di realtà per non cadere vittime delle loro allucinazioni.

24 giugno 2026 - Cassandra aveva già parlato del lavoro di Peter Steinberger, persona certo non priva di talento, essendo ad esempio l'autore di PSPDFKit, libreria oggi in esecuzione su una miliardata di dispositivi.

Aveva parlato di Steinberger perché questo signore ha recentemente scritto OpenClaw, la madre di tutti gli agenti; in seguito è entrato a far parte di OpenAI per lavorare, indovinate, sugli agenti.

Cassandra oggi se ne occupa nuovamente per la sua ultima dichiarazione, per non dire allucinazione. Steinberger ha infatti scritto che bisognerebbe smettere di dare istruzioni agli agenti di programmazione e iniziare a *«progettare loop che diano istruzioni ai propri agenti»*.
In questa "allucinazione" non è solo. Boris Cherny, che gestisce Claude Code presso Anthropic, ha dichiarato in pubblico che non dà più istruzioni a Claude. Ha dei *loop* in esecuzione che danno istruzioni a Claude e decidono cosa fare. «*Il mio lavoro è scrivere loop*», ha detto, definendo questa transizione come quella che tutti noi compiremo entro la fine dell'anno.

Per riassumere, come fa Ed Zitron nella sua ultima newsletter, la situazione descritta è "*dobbiamo smettere di dare comandi e iniziare a scrivere il sistema che li impartisce, salendo di un*

ulteriore gradino sulla scala dell'astrazione: schede perforate, linguaggio assembly, C, linguaggi di alto livello e ora i loop".

Un commento sotto il post di Steinberger porta il ragionamento alle solite estreme conseguenze: il passo successivo sarà un loop che progetta i loop, con il conseguente licenziamento del progettista perché diventato ormai un peso morto.

Questa è la direzione descritta dalla più recente allucinazione di moda sugli LLM, la stessa indicata dalle precedenti allucinazioni; creare disoccupati ed indurre stupidità.

In un mondo dove persino Cassandra ha generato del codice con l'ausilio di Claude, continuare a rilanciare sull'autogenerazione del codice da parte degli LLM è tuttavia l'unica strada utile per dimostrare l'indispensabilità degli LLM agli investitori.

Infatti la precedente *allucinazione dell'AGI*, l'intelligenza artificiale generale in grado di risolvere qualsiasi problema, compreso quello di automigliorarsi, non trova più molti sostenitori, in assenza di un qualsiasi indizio scientifico su come realizzarla.

Anche l'allucinazione dell'*autocoscienza degli LLM come comportamento emergente*, per quanto molto elegante provoca ormai solo scetticismo tra coloro che devono continuare a tirare fuori i soldi per sostenere la bolla degli LLM, a meno ovviamente che non facciano parte delle *comunità accelerazioniste*.

Appare quindi abbastanza evidente come i loop siano solamente l'ultimo argomento che i tech-bro (© DataKnightmare) utilizzano nella loro instancabile attività propagandista per tenere in piedi la bolla finanziaria.

Ma perché anche il boss di Nvidia Jensen Huang, alla testa dell'azienda più capitalizzata del mondo, nonché l'unica che sta facendo profitti "quasi reali" ("quasi" perché i suoi clienti/debitori sono destinati a fallire) sposa il nuovo concetto di *loop*, come ha entusiasticamente dichiarato?

E' abbastanza evidente; i loop sono un metodo per portare il consumo di token ad un livello ancora superiore, e l'unico prodotto di Nvidia sono appunto le GPU, le "*fabbriche di token*".

In tutto questo i tech-bro continuano ovviamente a ignorare il problema della sostenibilità di un uso crescente degli LLM, sostenibilità non dal punto di vista finanziario o delle conseguenze sociali, ma semplicemente da quello "materiale".

Risolvere problemi deterministici "bruciando" token è molto, ma davvero molto inefficiente da un punto di vista computazionale, e l'aumento continuo delle necessità di token, previsto da qui al 2030 per sostenere le necessità finanziare del circo degli LLM, è non solo insostenibile, ma semplicemente irrealizzabile, date le risorse materiali che richiederebbe.

I tech-bro oggi quantificano i datacenter di cui avranno bisogno utilizzando la ridicola *unità di misura* del loro consumo elettrico, perché non sanno definirli in altro modo.

Per cui, dovendoli riempire delle GPU di Nvidia, le più recenti delle quali consumano 1,5 Kilowatt l'una, cioè quanto un normale appartamento, favoleggiano di aver bisogno di datacenter da 1,2 Gigawatt, cioé della potenza elettrica prodotta da una grande centrale nucleare.

Ora, i datacenter più grandi esistenti sono "*grandi*" al massimo 100-300 Megawatt (ve la cavate bene con gli ordini di grandezza, vero?) e richiedono un minimo di tre anni dalla progettazione al completamento. In tutto il mondo ce ne è attualmente solo uno in costruzione, da 300 Megawatt.

Nessuno sa davvero come realizzare, nella pratica, un datacenter da un Gigawatt, qualunque cosa esso sia. Tutto questo dimostra, aldilà di ogni ragionevole dubbio, che da qui al 2029

nessun altro datacenter verrà realizzato, qualsiasi cosa i tech-bro dicano o pianifichino.

E questa in fondo è una buona cosa, perché passati i tre anni per costruirli non ci sarebbero comunque i Gigawatt di potenza elettrica per alimentarli. Costruire una centrale elettrica di grandi dimensioni richiede dai tre-quattro anni, se convenzionale, ai dieci-quindici anni se nucleare, e questo ipotizzando di realizzarla a rotta di collo senza il minimo inconveniente.

Anche qui vale la pena di notare che le centrali elettriche in fase di realizzazione in tutto il mondo bastano appena per sostenere l'aumento dei consumi convenzionali, e non sono certo pensate per alimentare i nuovi datacenter dei tech-bro; per quello ce ne vorrebbero ancora molte altre.

Ma anche questa in fondo, molto in fondo, è una buona cosa, perché se le centrali aggiuntive venissero realizzate non ci sarebbero comunque le reti elettriche in grado di trasmetterne la potenza ai datacenter; le reti elettriche sono strutture poco considerate ma enormi e complesse, ricche di parti vetuste, e che devono continuare a funzionare incessantemente, pena i blackout.

L'unica soluzione che possa operare con le reti elettriche attuali sarebbe quella di dotare ogni datacenter della sua centrale elettrica privata, possibilmente nucleare, nonché di un buon numero di gruppi elettrogeni diesel o gas per sopperire all'indisponibilità periodica della centrale per manutenzione o guasti.

I "*tre anni*", ricorrenti nelle considerazioni appena fatte, potrebbero sembrare "ragionevoli" per realizzare le prime risorse necessarie a mantenere la bolla degli LLM, ma così non è.

Dal punto di vista dei calcoli finanziari dei tech-bro, tre anni sono l'eternità, visto che le analisi di alcune nuove Cassandre stimano a non più di un anno da adesso lo scoppio della bolla, od almeno un drastico ridimensionamento di tutto il settore degli LLM, operato da chi vuole (o vorrebbe) rientrare dei circa mille miliardi investiti fino ad oggi.

Per riassumere il bagno di realtà che la vostra profetessa preferita vi ha oggi imposto, il concetto di *loop* costituisce un nuovo tentativo di far consumare un numero sempre crescente di token, per poter "sostenere" la pianificazione finanziaria del settore degli LLM, consentendo sulla carta di "predire" ritorni economici crescenti per le enormi spese in conto capitale (CAPEX) principalmente per nuovi datacenter, già programmate per il futuro.

Spese sempre crescenti necessarie per mantenere in vita aziende come OpenAI od Anthropic che, se fossero anche solo leggermente "normali", dovrebbero essere fallite già da tempo.

Questi spese in conto capitale non potranno mai essere realmente fatte, data l'impossibilità di realizzare le infrastrutture, e soprattutto di farlo nei tempi necessari a non far crollare la baracca finanziaria.

E quindi?

E quindi vivremo necessariamente in tempi interessanti.

Scrivere a Cassandra - Mastodon - Buttondown.com
Videorubrica "Quattro chiacchiere con Cassandra
Lo Slog (Static Blog) di Cassandra
L'archivio di Cassandra: scuola, formazione e pensiero

Cassandra per i posteri: l'archivio su Internet Archive

Licenza d'utilizzo: i contenuti di questo articolo, dove non diversamente indicato, sono sotto licenza Creative Commons Attribuzione—Condividi allo stesso modo 4.0 Internazionale (CC

BY-SA 4.0), tutte le informazioni di utilizzo del materiale sono disponibili a questo link.